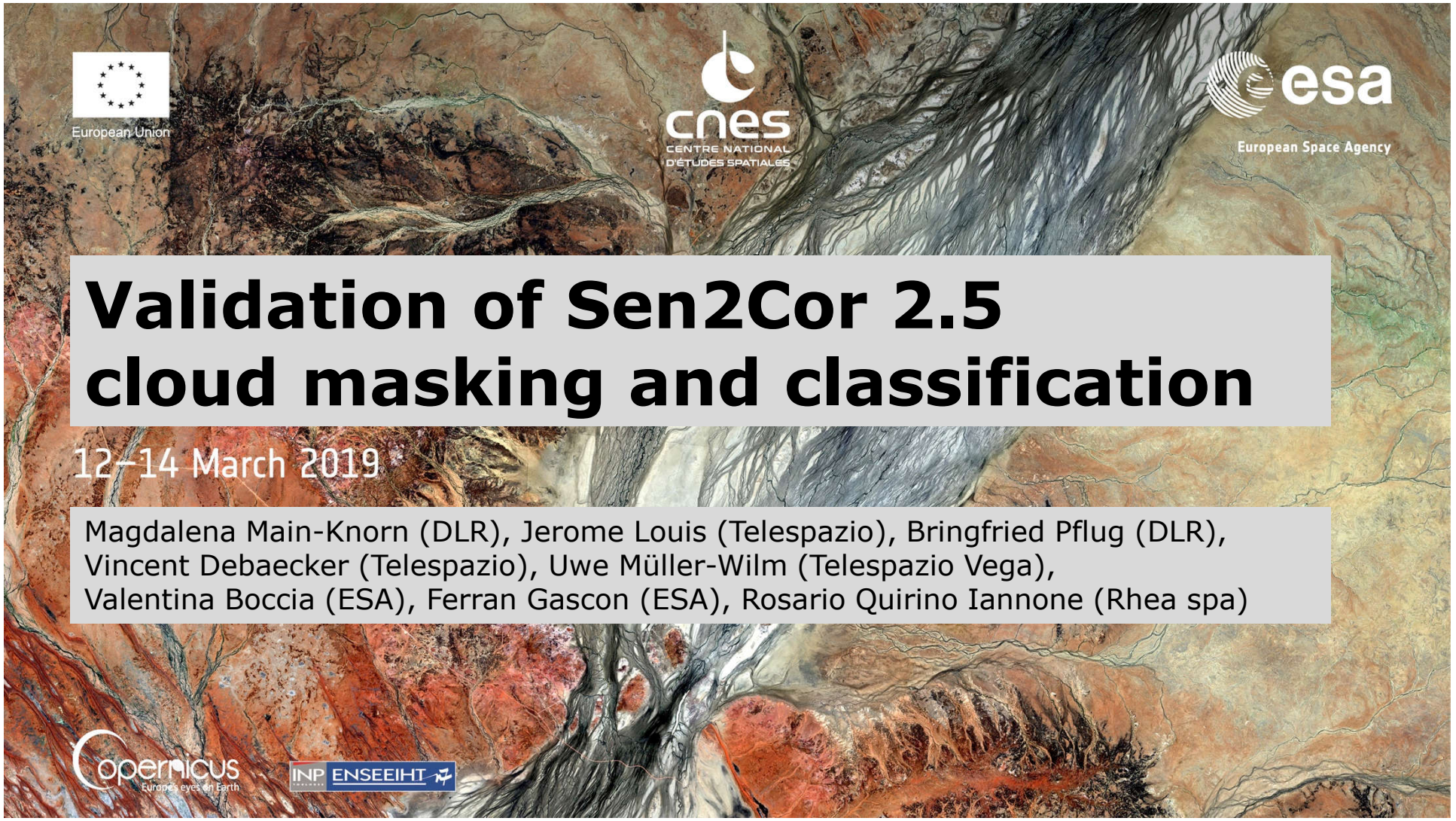

European Union
cnes
CENTRE NATIONAL D'ÉTUDES SPATIALES
esa
European Space Agency
Validation of Sen2Cor 2.5 cloud masking and classification
12–14 March 2019
Magdalena Main-Knorn (DLR), Jerome Louis (Telespazio), Bringfried Pflug (DLR),
Vincent Debaecker (Telespazio), Uwe Müller-Wilm (Telespazio Vega),
Valentina Boccia (ESA), Ferran Gascon (ESA), Rosario Quirino Iannone (Rhea spa)
opernicus
Europe's eyes on Earth
INP ENSEEIHT

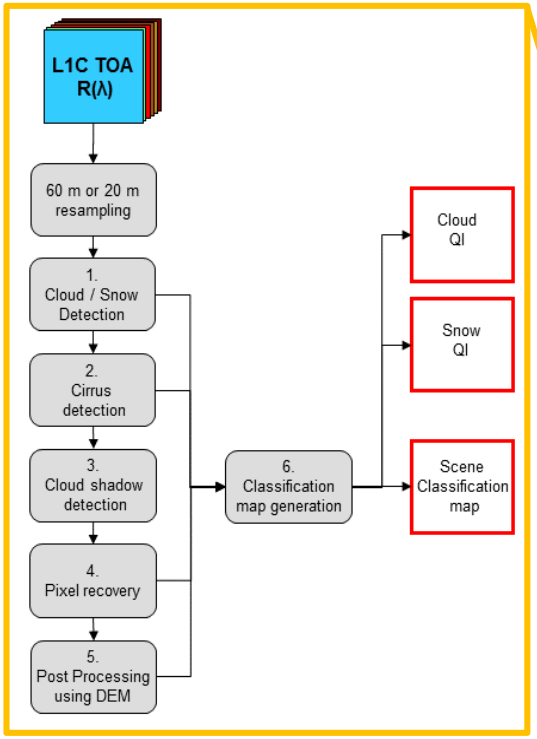
Outline

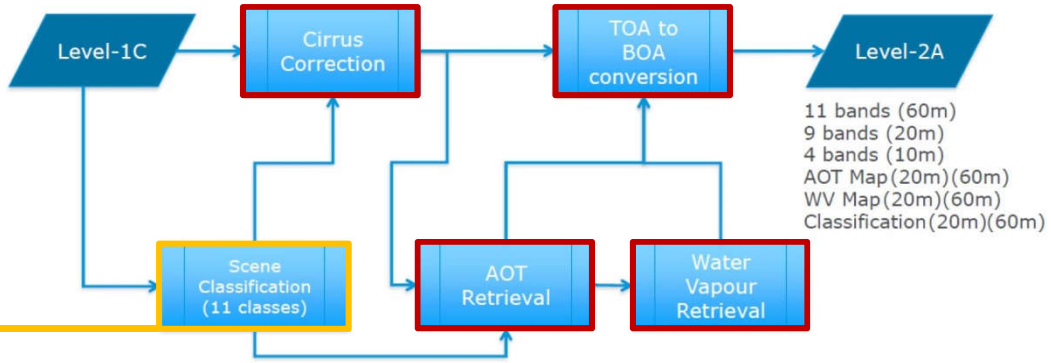
1. Sen2Cor Processor
2. Validation Dataset
3. Validation Procedure
4. Results and Examples
5. Conclusions
6. Outlook

# 1. Sen2Cor 2.5 Processor

- Atmospheric correction processor, developed by Telespazio on behalf of ESA
- Mono-temporal processor: corrects single-date Sentinel-2 L1C Top-Of-Atmosphere (TOA) products from the effects of the atmosphere and delivers a L2A Bottom-Of-Atmosphere (BOA) surface reflectance product
- Additional outputs: Aerosol Optical Thickness (AOT) map, Water Vapour (WV) map and Scene Classification (SCL) map with Quality Indicators for cloud and snow probabilities, in JPEG 2000 image format
- L2A data production run on the global systematic basis with dissemination through the Copernicus Open Access Hub since December 2018
- Processor can be freely downloaded (latest release 2.5) and used for processing L1C images

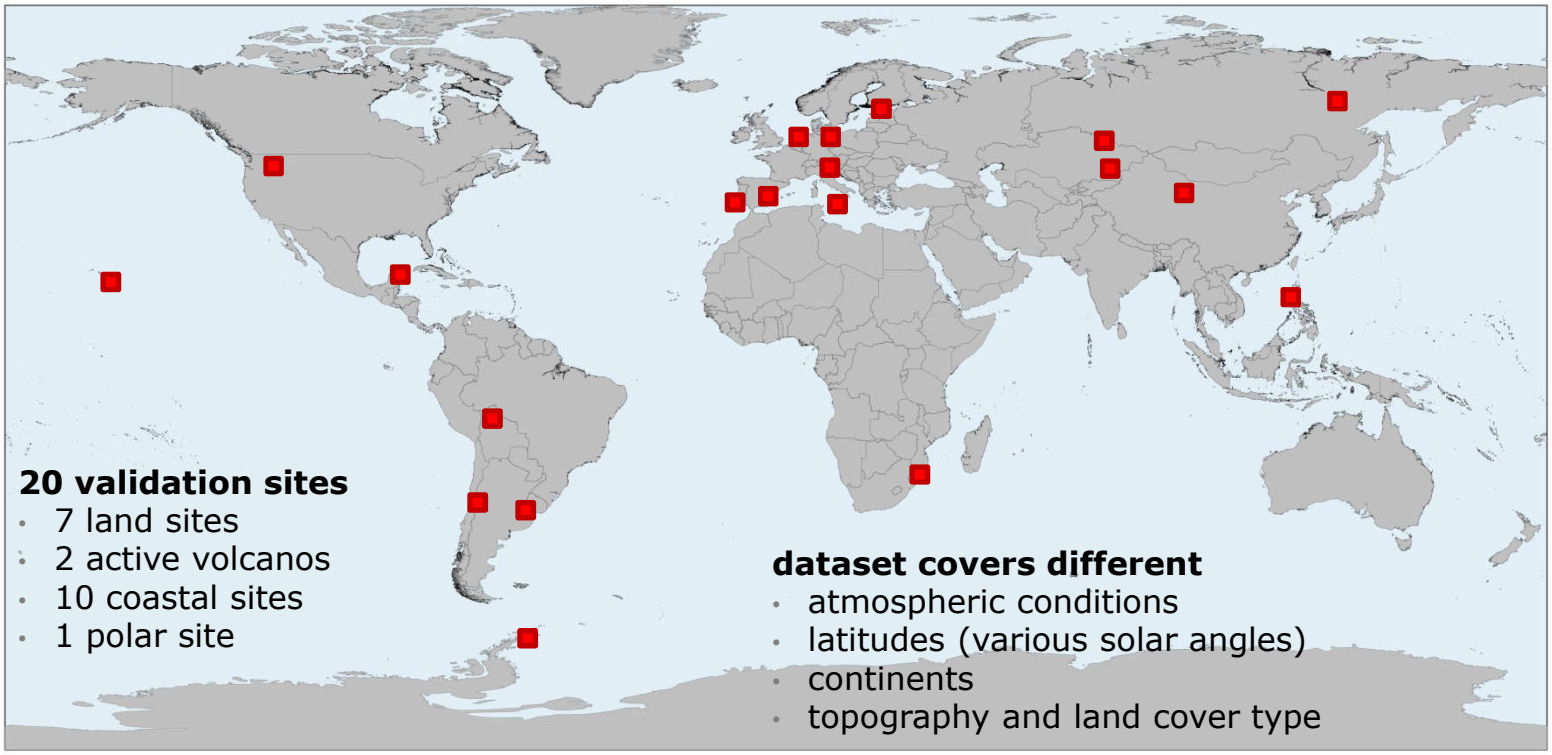
# 1. Sen2Cor 2.5 Processor Framework - Scene Classification Module (SCL)




- Two main modules: Scene Classification (SCL), Atmospheric Correction (AC)
- DEM downloaded automatically by the processor (SRTM) or provided by the user in DTED format
- SCL is mainly based on a series of threshold tests on L1C spectral bands, band ratios and indices
- ESA CCI auxiliary data support for water and snow detection

# 2. Validation Dataset



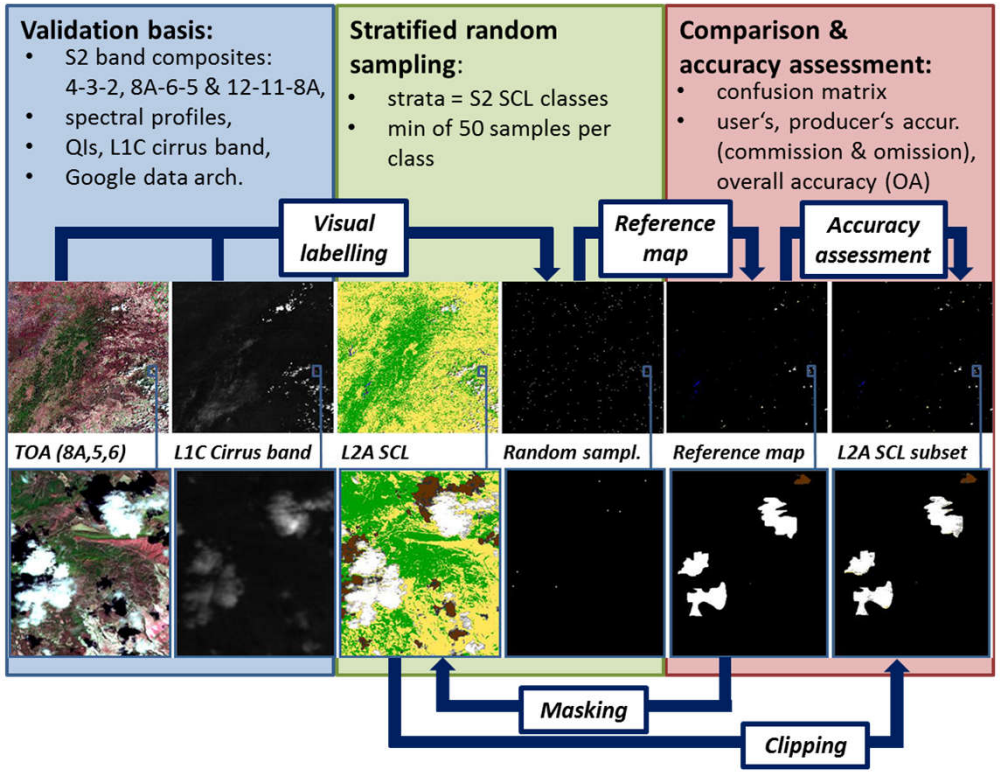
**20 validation sites**
- 7 land sites
- 2 active volcanos
- 10 coastal sites
- 1 polar site

**dataset covers different**
- atmospheric conditions
- latitudes (various solar angles)
- continents
- topography and land cover type

# 2. Validation Dataset

| Country | Site | Tile | Date | L1C clouds | Land cover | Special features |
|---|---|---|---|---|---|---|
| Antarctic | Antarctic | 21EVK | 04.02.2016 | 9,91 | snow, ice, water | clouds over snow, clouds and shadows over water |
| Argentina | Casleo | 19HDE | 12.08.2016 | 21,73 | snow, rural, mountains | transparent clouds, cloud shadows, topo shadows |
| Argentina | Buenos Aires | 21HUC | 27.08.2018 | 0,00 | urban, rural, water | sediment in water, very bright objects |
| Bolivia | Puerto Siles | 19LHF | 06.09.2018 | 0,00 | rural, vegetation, water | different water color and sediment, bright soil |
| China | Dunhuang | 46TFK | 22.01.2018 | 24,14 | snow, rural, desert | no clouds, solar panels, |
| Estonia | Tallin | 35VLG | 14.07.2018 | 2,09 | urban, rural, water | algal bloom, bright objects |
| Germany | Berlin | 33UUU | 04.05.2018 | 0,87 | urban, rural, water | asparagus fields |
| Italy | Etna Volcano | 33SVB | 09.03.2017 | 6,86 | snow, rural, urban, water | active volcano, burning and burned area, ash cloud |
| Italy | Venice | 32TQR | 16.09.2018 | 2,42 | coast, industry, urban, rural, water | black coastline, topographic shadows, bright object |
| Kazakhstan | Pavlodar | 43UET | 26.07.2018 | 0,13 | rural, industrial and natural water | different water colours, |
| Kazakhstan | Balkhash | 43TFM | 30.07.2018 | 7,18 | water, rural | black coastal line |
| Mexico | Cancun | 16QDJ | 27.05.2018 | 6,68 | rural, vegetation, water | waves, sediment, algas(?), clouds |
| Mosambique | Maputo | 36JVS | 13.07.2018 | 0,00 | rural, urban, water | burned area, sediment in water, |
| Netherlands | Amsterdam | 31UFU | 13.09.2018 | 4,73 | rural, urban, vegetation, water | black coastline, clouds & cloud shadows over water |
| Phillipines | Manila | 51PTS | 19.03.2018 | 1,44 | urban, rural, water | sediment / aquacultures, very bright objects, |
| Portugal | Lagos | 29SNB | 08.08.2018 | 0,00 | rural, water | waves, sediment, burned area, bright objects |
| Russia | Yakutsk | 52VEP | 08.03.2016 | 61,53 | snow, rural | transparent clouds |
| Spain | Barrax | 30SWH | 09.05.2017 | 17,67 | rural, vegetation, mountains | cumulus and cirrus clouds |
| Spain | Barrax | 30SWH | 19.05.2017 | 1,64 | rural, vegetation, mountains | cumulus clouds |
| USA | Rimrock | 11TMM | 12.05.2018 | 0,74 | snow, rural, urban, water | very bright objects |
| USA-Hawaii | Kilauea Volcano | 05QKB | 23.04.2018 | 28,41 | rural, water | active volcano, burning and burned area, ash cloud |

# 3. Validation Procedure




Statistical metrics:

**Overall accuracy (OA)**

*Percentage of correctly classified pixels.*

**Omission error (OE) ; Producer Accuracy (PA)**

*The share of reference pixels in that class that have been "omitted" in the classification image.*

*OE [%] = 100 - Producer Accuracy*

**Commission error (CE) ; User Accuracy (UA)**

*Percentage of class pixels in the classification image which are falsely classified.*

*CE [%] = 100 - User Accuracy*

Statistics are calculated for:

- All classes (original SCL product)
- Valid pixels (over land and water) separation
- Clouds (medium- and high probability clouds, cirrus) separation

# 3. Validation Procedure - Limitations

- Stratified random sampling is based on classification SCL (class = strata) - it can bias validation results, where Sen2Cor classification failed
- Reference samples set is inbalanced, semi-proportional to the class area
- Definition of thin cirrus and visual interpretation of clouds is subjective & challenging
- Internal test on the subjectivity of the validation method (4 validating persons, 2 products) revealed quite stable results (st.dev of classification accuracy ~5-6%). More extended tests are required.
- In preparation classification validation protocol for L2A to assure validation quality and comparability between algorithms

# 4. Results – per site

| Country | Site | Tile | Date | L1C clouds | OA | OA valid | OA clouds | Pixel validated |
|---|---|---|---|---|---|---|---|---|
| Antarctic | **Antarctic** | 21EVK | 04.02.2016 | 9,91 | 94,7 | 96,8 | 98,8 | 527803 |
| Argentina | **Casleo** | 19HDE | 12.08.2016 | 21,73 | 63,8 | 86,1 | 98,1 | 186238 |
| Argentina | **Buenos Aires** | 21HUC | 27.08.2018 | 0,00 | 91,8 | 97,3 | no clouds | 31841 |
| Bolivia | **Puerto Siles** | 19LHF | 06.09.2018 | 0,00 | 91,1 | 94,9 | 94,8 | 80580 |
| China | **Dunhuang** | 46TFK | 22.01.2018 | 24,14 | 57,3 | 66,2 | no clouds | 105454 |
| Estonia | **Tallin** | 35VLG | 14.07.2018 | 2,09 | 84,3 | 90,4 | 95,6 | 71773 |
| Germany | **Berlin** | 33UUU | 04.05.2018 | 0,87 | 93,4 | 96,5 | no clouds | 51964 |
| Italy | **Etna Volcano** | 33SVB | 09.03.2017 | 6,86 | 95,8 | 97,9 | 99,4 | 132340 |
| Italy | **Venice** | 32TQR | 16.09.2018 | 2,42 | 80,3 | 91,0 | 89,0 | 103780 |
| Kazakhstan | **Pavlodar** | 43UET | 26.07.2018 | 0,13 | 70,1 | 82,6 | 83,2 | 90803 |
| Kazakhstan | **Balkhash** | 43TFM | 30.07.2018 | 7,18 | 77,4 | 93,7 | 95,7 | 94578 |
| Mexico | **Cancun** | 16QDJ | 27.05.2018 | 6,68 | 90,5 | 93,8 | 95,2 | 32082 |
| Mosambique | **Maputo** | 36JVS | 13.07.2018 | 0,00 | 80,5 | 85,1 | 82,5 | 90043 |
| Netherlands | **Amsterdam** | 31UFU | 13.09.2018 | 4,73 | 71,1 | 81,5 | 88,7 | 100018 |
| Phillipines | **Manila** | 51PTS | 19.03.2018 | 1,44 | 82,1 | 90,0 | 91,6 | 106263 |
| Portugal | **Lagos** | 29SNB | 08.08.2018 | 0,00 | 96,8 | 97,3 | no clouds | 69753 |
| Russia | **Yakutsk** | 52VEP | 08.03.2016 | 61,53 | 69,9 | 93,8 | 92,9 | 177983 |
| Spain | **Barrax** | 30SWH | 09.05.2017 | 17,67 | 64,6 | 96,9 | 98,7 | 141546 |
| Spain | **Barrax** | 30SWH | 19.05.2017 | 1,64 | 90,5 | 98,7 | 99,5 | 104799 |
| USA | **Rimrock** | 11TMM | 12.05.2018 | 0,74 | 90,2 | 98,2 | 99,2 | 103394 |
| USA-Hawaii | **Kilauea Volcano** | 05QKB | 23.04.2018 | 28,41 | 60,4 | 75,4 | 74,2 | 118357 |
| | | | | **Average OA:** | **80,9** | **Averaged ref per SCL:** | | **120066** |
| | | | | | | **Total validated pixels:** | | **2521392** |

**Best case:**
**OA > 90%;**
**OA clear pixels > 94%**
**OA clouds > 95%**

**Moderate case:**
**OA 70-85%;**
**OA clear pixels 80-94%**
**OA clouds 80-96%**

**Poor case:**
**OA >= 64 %**
**but**
**OA clear pixels > 86%**
**OA clouds > 93%**

**The worst case:**
**OA <= 60%;**
**OA clear pixels &**
**OA clouds < 75%**

# 4. Results – per valid and cloudy pixels

| Valid pixels over land and water | | | | |
|---|---|---|---|---|
| | **Land-Water** | **Others** | **Sum** | **Commission error** |
| **Valid pixels Land-Water** | 1119112 | 137273 | 1256385 | **10.9** |
| **Others** | 83033 | 1181974 | 1265007 | **6.6** |
| **Sum** | 1202145 | 1319247 | 2521392 | |
| **Omission error** | **6.9** | **10.4** | | |
| **All clouds** | | | | |
| | **Clouds** | **Others** | **Sum** | **Commission error** |
| **Clouds** | 577133 | 34464 | 611597 | **5.6** |
| **Others** | 114256 | 1795539 | 1909795 | **6.0** |
| **Sum** | 691389 | 1830003 | 2521392 | |
| **Omission error** | **16.5** | **1.9** | | |

(+) Good performance for valid pixels detection (CE=10,9%; OE=6,9%)

(–) High omission error for clouds detection (CE=5,6%; OE=16,5%)

- Increased CE for valid pixels means that too many invalid pixels (mainy clouds) were classified as valid
- High OE for clouds means that too many clouds were classified as valid pixels (mainly on the cloud borders)
- At this Sen2Cor evolution stage dilation of cloud mask would not decrease cloud omission error !!

# 4. Results – per selected classes

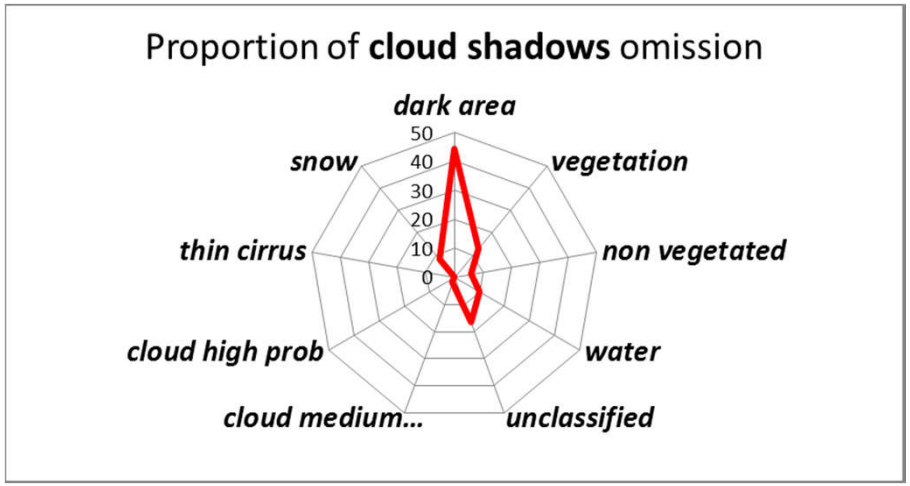
| Class | Commission error | Omission error |
|---|---|---|
| **Cloud shadows** | 4,5 | 55,6 |
| **Vegetation** | 17,7 | 0,5 |
| **Non vegetated** | 16,1 | 16,7 |
| **Water** | 5,6 | 1,9 |
| **Snow** | 5,5 | 10,3 |

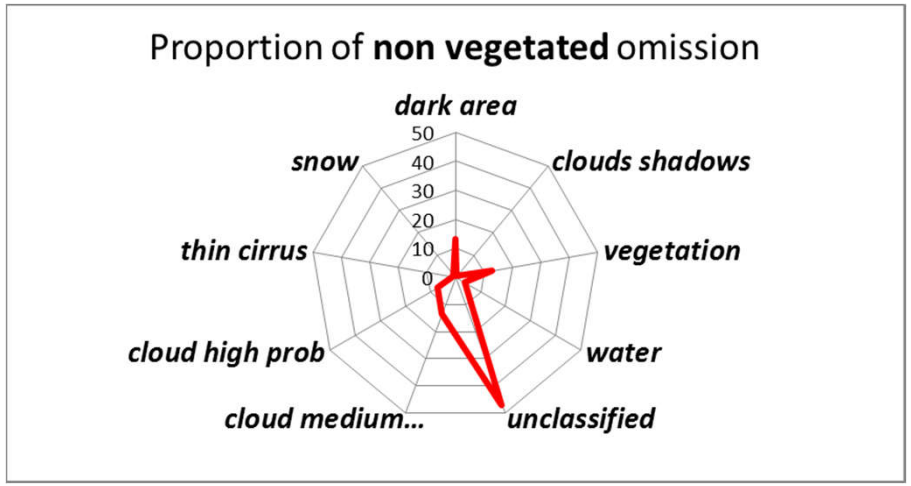
(+) Very low omission for vegetation and water

(+) Very low commission for cloud shadows, water and snow/ice

(–) Very high omission for cloud shadows

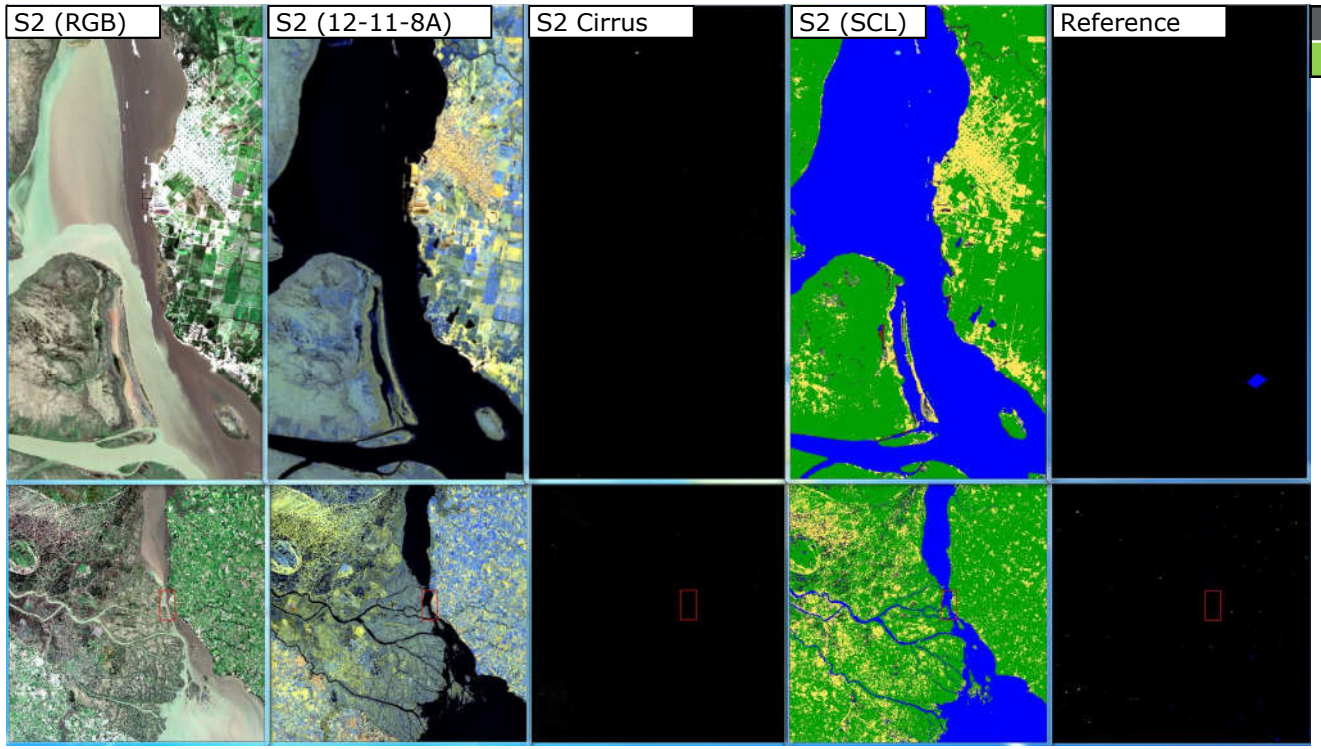
(–) Higher omission for non-vegetated pixels



# 4. Results/ Example 1: Buenos Aires

## – urban area, sediment in water




| OA | OA valid | OA clouds |
|---|---|---|
| 91,8 | 97,3 | no clouds |

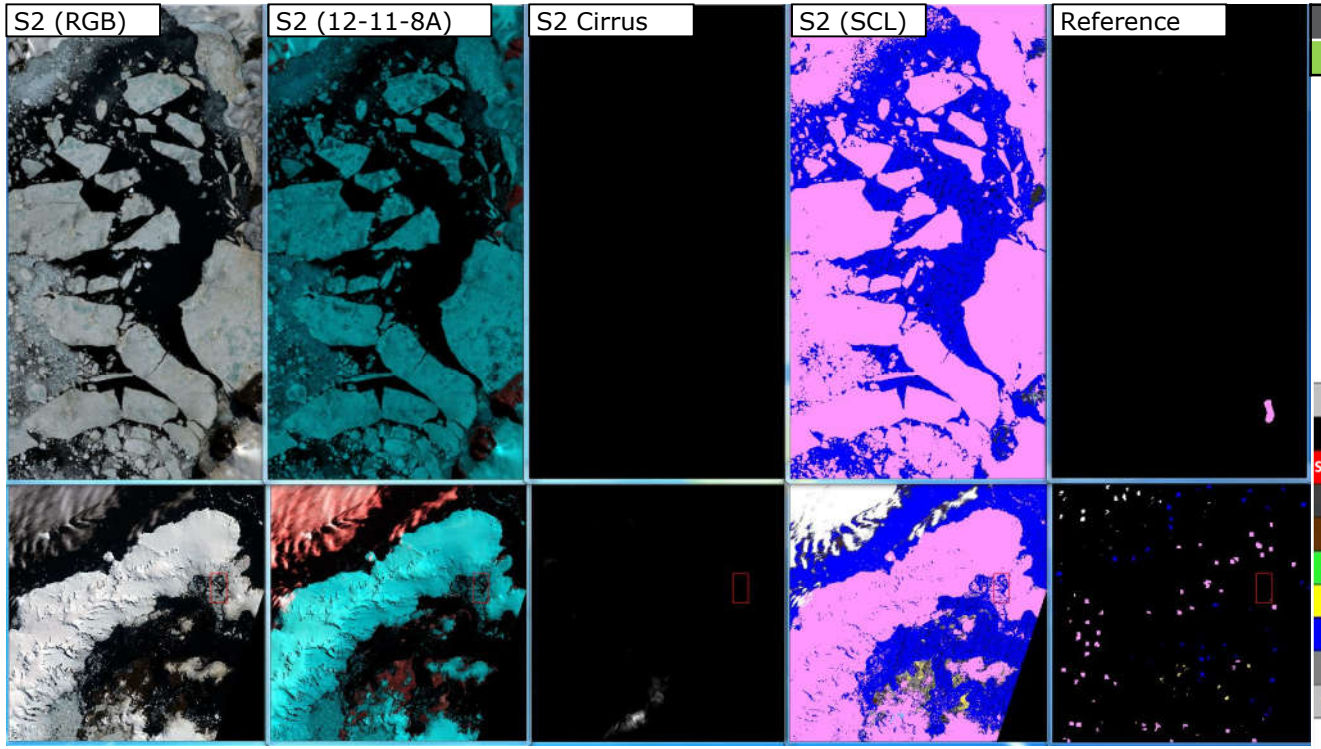
(+) Different water colors and ship traffic possible to identify (but it is not the focus of SCL classification)

(+) Water, vegetation and non vegetated recognized correctly (also on bright urban surfaces)

| Class |
|---|
| NO_DATA |
| SATURATED_or_DEFECTIVE |
| DARK_AREA_PIXELS |
| CLOUD_SHADOWS |
| VEGETATION |
| NON_VEGETATED |
| WATER |
| UNCLASSIFIED |
| CLOUD_MEDIUM_PROBAB |
| CLOUD_HIGH_PROBAB |
| THIN_CIRRUS |
| SNOW |

# 4. Results/ Example 2: Antarctic

## – snow, ice, clouds



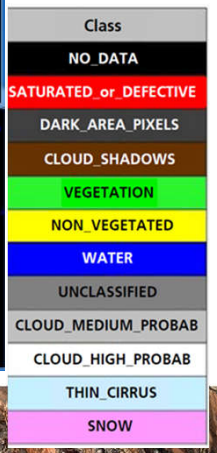

S2A_USER_PRD_MSIL2A_PDMC_20160206T015135_R095_V20160204T131909_20160204T131909.SAFE

| OA | OA valid | OA clouds |
|---|---|---|
| 94,7 | 96,8 | 98,8 |

(+) Clouds over water and snow classified properly

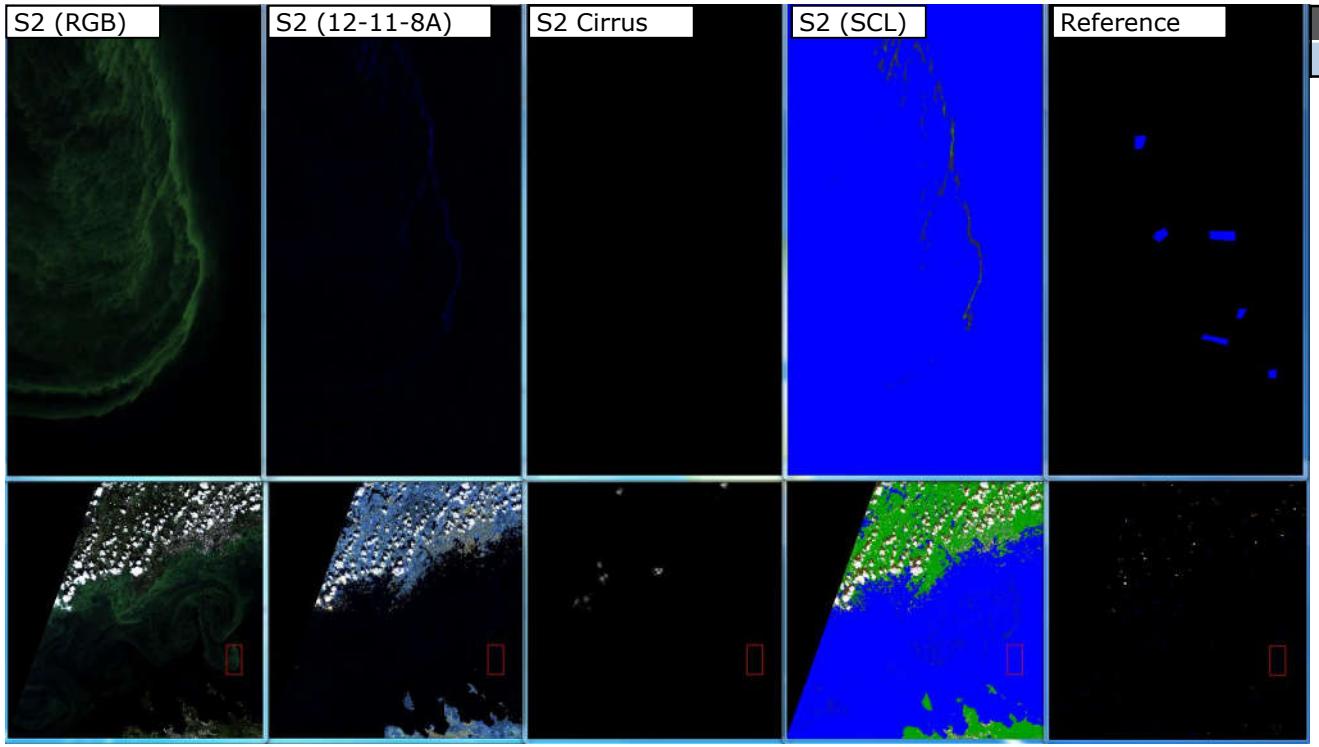
(+) Water, snow and ice classified correctly

(–) Omission of non vegetated towards clouds



# 4. Results/ Example 3: Tallin

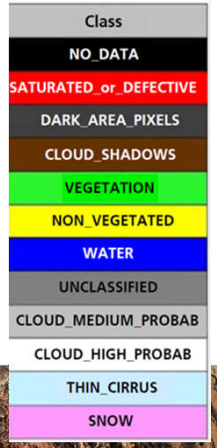
## – algal bloom, urban




| OA | OA valid | OA clouds |
|---|---|---|
| 84,3 | 90,4 | 95,6 |

(+) Cloud, vegetation and water classified properly

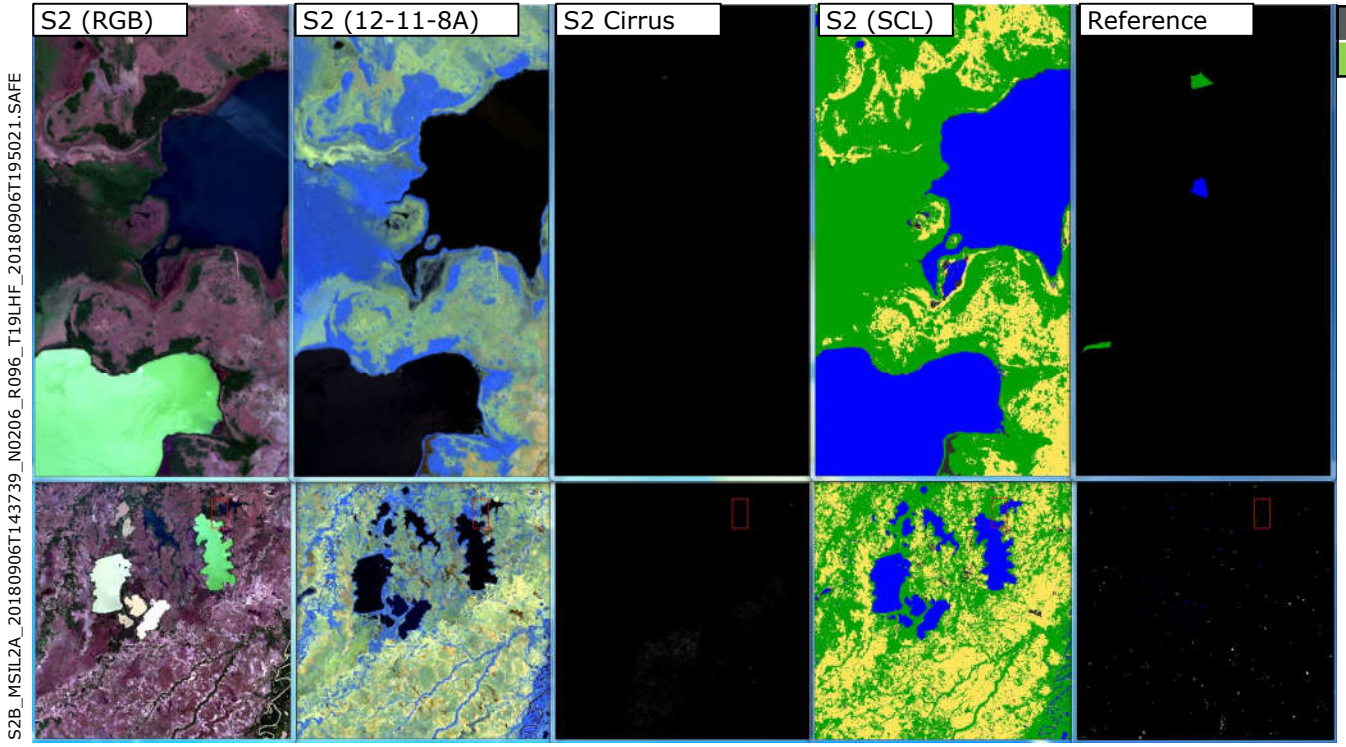
(+) Algal bloom possible to identify (but it is not the focus of SCL classification)




(–) Lower OA of clear pixels due to misclassification of bright harbor area as clouds

# 4. Results/ Example 4: Puerto Siles

## – water colors, sunglint




| OA | OA valid | OA clouds |
|---|---|---|
| 91,1 | 94,9 | 94,8 |

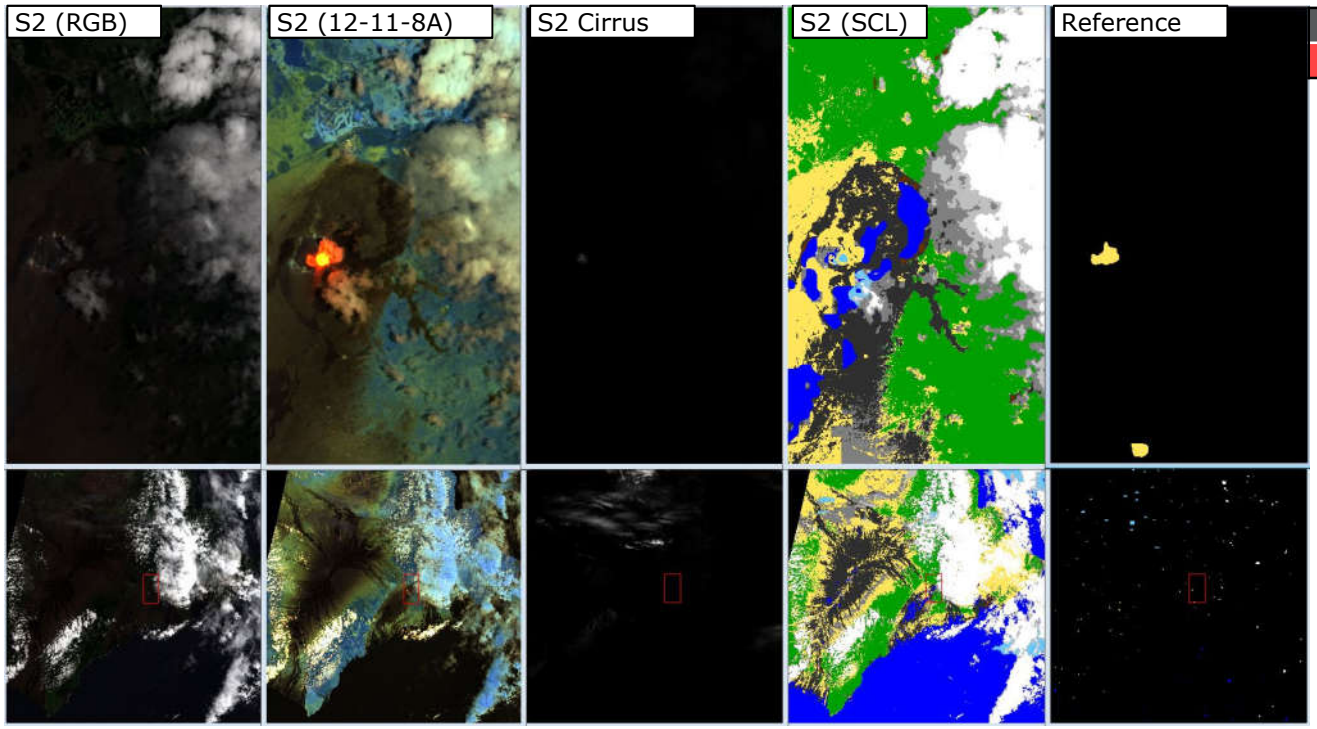
(+) Water classified properly despite of different water colors

(+) Vegetation and non vegetated recognized correctly

| Class |
|---|
| NO_DATA |
| SATURATED_or_DEFECTIVE |
| DARK_AREA_PIXELS |
| CLOUD_SHADOWS |
| VEGETATION |
| NON_VEGETATED |
| WATER |
| UNCLASSIFIED |
| CLOUD_MEDIUM_PROBAB |
| CLOUD_HIGH_PROBAB |
| THIN_CIRRUS |
| SNOW |

# 4. Results/ Example 5: Kilauea Volcano

## – fire, burned area, clouds, smoke




| OA | OA valid | OA clouds |
|---|---|---|
| 60,4 | 75,4 | 74,2 |

(+) Cloud, vegetation & non vegetation classified properly

(+) Fire and burned area partially recognized; burned area classified as dark area

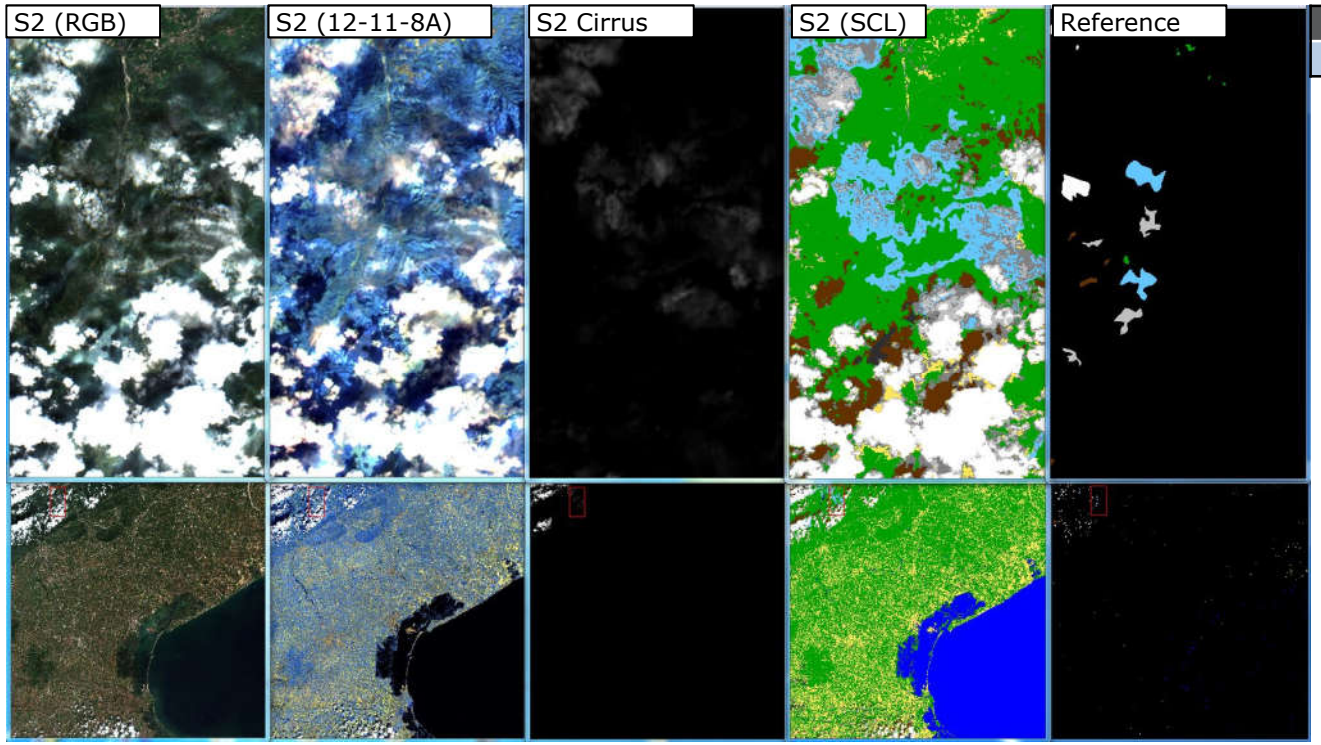
| Class |
|---|
| NO_DATA |
| SATURATED_or_DEFECTIVE |
| DARK_AREA_PIXELS |
| CLOUD_SHADOWS |
| VEGETATION |
| NON_VEGETATED |
| WATER |
| UNCLASSIFIED |
| CLOUD_MEDIUM_PROBAB |
| CLOUD_HIGH_PROBAB |
| THIN_CIRRUS |
| SNOW |

(–) Commission of water around volcano

(–) Unclear treatment of haze or smoke (no class defined)

# 4. Results/ Example 6: Venice

## – clouds and cirrus confusion




| OA | OA valid | OA clouds |
|---|---|---|
| 80,3 | 91,0 | 89,0 |

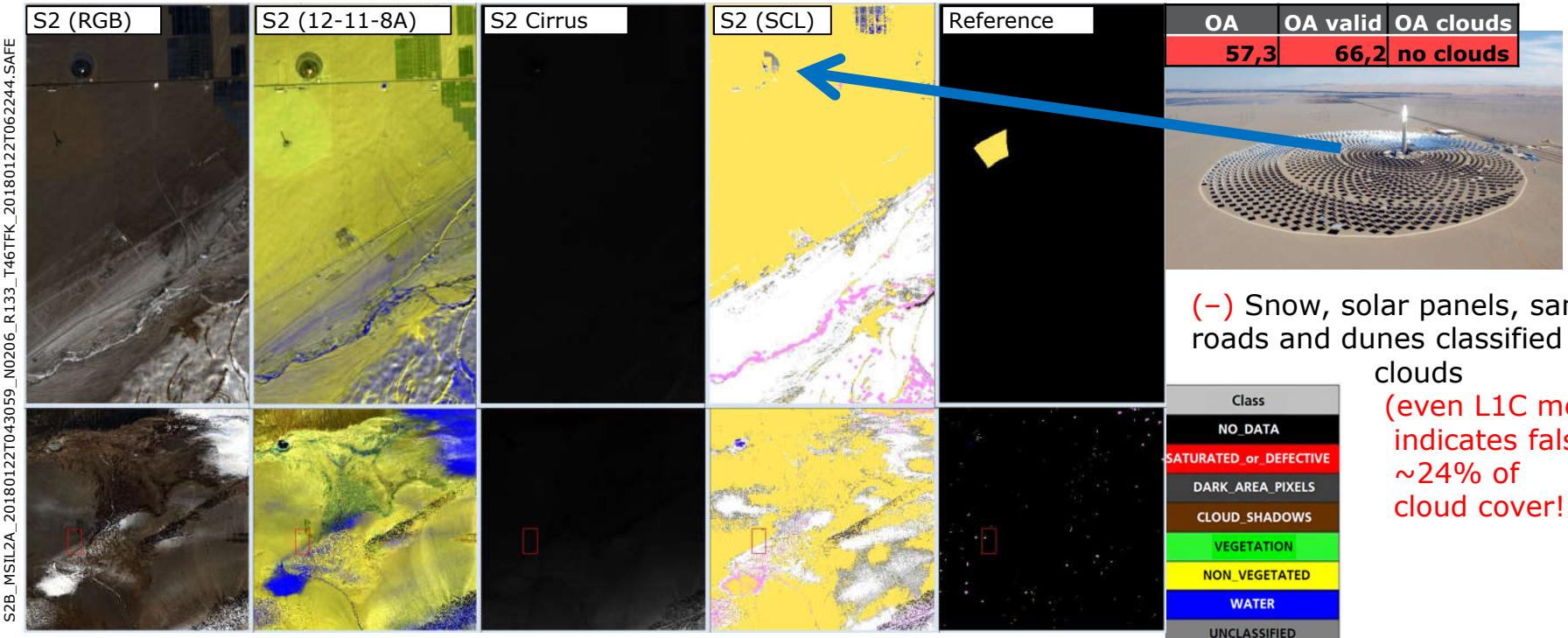
(+) Water, vegetation & non vegetated classified properly

(–) Confusion between cloud classes

| Class |
|---|
| NO_DATA |
| SATURATED_or_DEFECTIVE |
| DARK_AREA_PIXELS |
| CLOUD_SHADOWS |
| VEGETATION |
| NON_VEGETATED |
| WATER |
| UNCLASSIFIED |
| CLOUD_MEDIUM_PROBAB |
| CLOUD_HIGH_PROBAB |
| THIN_CIRRUS |
| SNOW |

# 4. Results/ Example 7: Dunhuang

## – snow, very bright sand, solar panels

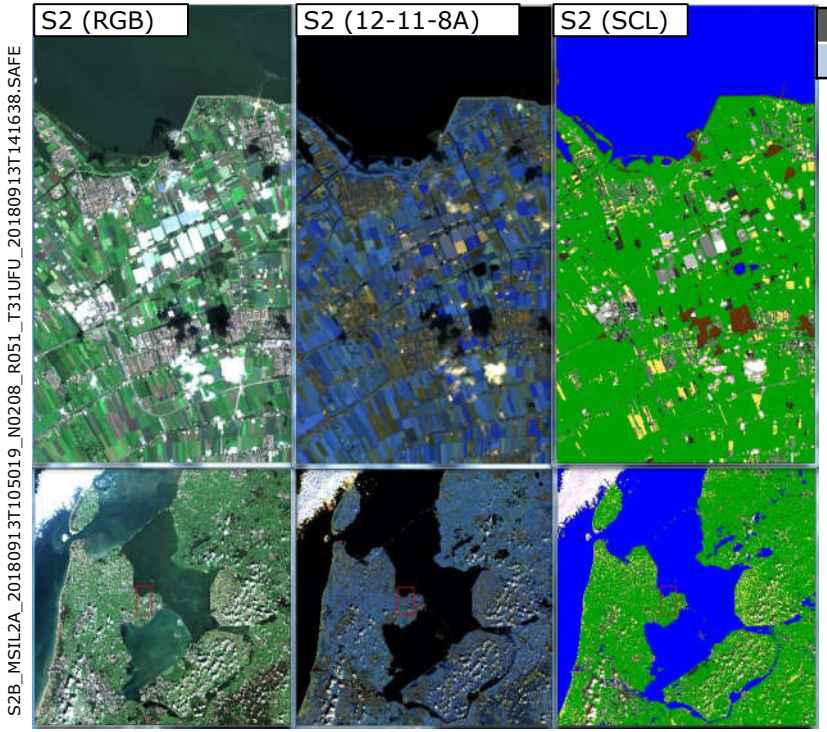



(–) Snow, solar panels, sandy roads and dunes classified as clouds

(even L1C meta indicates false ~24% of cloud cover!!!)

# 4. Results/ Example 8: Amsterdam

## – agriculture, bright and dark urban objects




| OA | OA valid | OA clouds |
|---|---|---|
| 71,1 | 81,5 | 88,7 |

(-) High misclassification rate for bright objects e.g. urban areas, highly reflecting roofs, asparagus fields covered by foil – classified as clouds

(-) High omission and commision errors of medium cloud probability and cirrus classes

– Confusion between cloud classes definitions and some overlapping problems

(-) High omission and commision errors for dark area pixels

– Confusion between cloud shadows, topographic shadows and dark area pixel

– Confusion in the definition of dark area pixel (topographic shadow? also dark soil, burned area?)

# 5. Conclusions

- Very good separation of land, coastal line and water, even if affected by e.g. algae bloom or sediment
- High identification potential for burned areas – even if not a scope of SCL classification
- Averaged overall performance of Sen2Cor 2.5 for 21 products reaches 80,9%
  - Very high accuracy for water, vegetation and snow / ice classes
  - Very low commission error for cloud shadows
  - High misclassification rate for bright objects (urban, arid) – improvement needed, work ongoing
- Good performance for valid pixels (over land and water) and clouds (3 cloud classes) detection:
  - CE and OE of clear pixels separation CE=10,9% and OE=6,9% respectively
  - CE and OE of cloud separation CE=5,6% and OE=16,5% respectively
  - Consolidated cloud mask could provide high quality support for the users applications – work ongoing for forthcoming Sen2Cor processor evolutions
- Sen2Cor processor can be freely downloaded and used http://step.esa.int/main/third-party-plugins-2/sen2cor/
- L2A Product Performance now reported in the L2A Data Quality Report:

https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2-L2A-Data-Quality-Report

# 6. Outlook

I priority:

- Improvement of SCL over bright surfaces to reduce the number of misclassification on urban area, arid and semi-arid sites towards clouds
- Improvement of SCL for cirrus detection in mountanious areas

} work ongoing

- Development of the Consolidated Cloud Mask (CCM)
- Evaluation o the Cloud Mask Dilation Strategy

II priority:

- Separation of cloud shadows and topographic shadows from other dark features allocated in the dark area class is aimed

- Building a representative independent reference data set is a huge work - validation protocol for L2A to assure validation quality needed.

# Many thanks for your attention

Magdalena Main-Knorn (DLR), Jerome Louis (Telespazio), Bringfried Pflug (DLR),
Vincent Debaecker (Telespazio), Uwe Müller-Wilm (Telespazio Vega),
Valentina Boccia (ESA), Ferran Gascon (ESA), Rosario Quirino Iannone (Rhea spa)

magdalena.main-knorn@dlr.de; www.DLR.de/eoc